

***Reviewer #1: Review of
Deriving Sediment Interstitial Water Remediation Goals (IWRGs) at Superfund
Sites for the Protection of Benthic Organisms from Direct Toxicity***

Charge Questions:

- (1) Is the document written in a style that will be accessible for users with a range of educational and technical backgrounds? **RESPONSE:** The document is written in a manner that will be accessible for the range of practitioners engaged at Superfund sites.
- (2) Is the described methodology sufficiently clear to be performed by Superfund remediation project managers, risk assessors, and consultants for Superfund sites? If not, please provide suggestions on how clarity can be improved. **RESPONSE:** The description of the methodology is understandable. However, there are several places in the document (noted in my specific comments) where additional detail will be required in order for practitioners to implement the guidance in a reliable and consistent manner.
- (3) Is the document missing any important concepts, sections, definitions, and/or text that should be provided in order to make the methodology truly implementable? **RESPONSE:** As noted above, additional detail and specificity is required in several places that are identified in my specific comments. I recommend in my specific comments that additional thought and development be devoted to considering application of the general approach to bioaccumulative compounds. I also recommend that Section 5 be recast into a description of a multiple lines-of-evidence framework for using the methodology to support decision making.
- (4) Are the illustrative examples for determining IWRGs complete enough to demonstrate how the IWRGs are derived? **RESPONSE:** It would be very useful to practitioners if specific case example applications, even if hypothetical, could be included in an appendix to the document as a way of illustrating the application of the methodology for different decision-making scenarios and/or management applications.
- (5) Is the methodology for deriving interstitial water remediation goals scientifically defensible? **RESPONSE:** The general methodology is defensible. However, proper application of the methodology will be key, especially in view of the variability in conditions at many Superfund sites and the assumptions and uncertainties associated with the method.
- (6) In implementing the methodology, site-specific K_{oc} s are used to convert the IWRGs on concentration basis in sediment interstitial water ($\mu\text{g/L}$) to concentrations in bulk sediment ($\mu\text{g/kg}$ dry weight). Is the discussion of the K_{oc} s adequate? Is the discussion of the conversion from concentrations in interstitial water to bulk sediment adequate? Is the discussion of which K_{oc} s should be used in the conversions adequate? **RESPONSE:** As related in my specific comments, I believe additional work is needed in regard to application of the methodology given the varying conditions that can be expected across many Superfund sites in terms of total

organic carbon content and the nature of that carbon in respect to partitioning. This is one of the most critical issues that will affect the quality of applications.

- (7) Passive sampling can be performed on any number of samples from a site; for example, on all samples where contaminants are measured in bulk sediment, on only the surface sediments, on the top and bottom of sediments cores, on the top and at the dredge depth of the sediments cores, on surface sediment and based on BAZ (biological active zone), or some other arrangement. Currently, the methodology allows flexibility (makes no recommendation) on which samples are measured using the passive sampling technique and how those data are used in the conversion from interstitial water IWRGs to bulk sediment IWRGs. The extremes in this process are a) perform one passive sampling measurement and assume all sediments are the same across the location of interest (horizontally and with depth) or b) perform passive sampling on all samples and develop 3-D contour plots with depth based upon concentrations in the interstitial water. Should the methodology make a recommendation on this issue? If so, provide your recommendation. **RESPONSE:** Rather than including only one approach to this within the guidance, I recommend that additional guidance be prepared in the form of a logical framework (including a conceptual model) that project teams could follow in determining how to best apply the methodology for their site conditions and decision-making/management needs.

- (8) Section 5 provides information on comparing toxicity test results and developed IWRGs. Is this section sufficiently clear for the non-experts in toxicity testing and/or passive sampling?

RESPONSE: See response to question 3.

Specific Comments:

1. The name IWRG is “conceptually” awkward, as we don’t remediate interstitial water. What is being proposed is really a Benthic Bioavailability Remediation Goal, or something similar.
2. Page 5. There are two separate issues in this part of the discussion: uncertainty and cost of application. These topics/themes should be highlighted more directly.
3. Page 5 past paragraph. Organic carbon doesn’t “cause” toxicity.
4. Page 9. Do you mean to exclude P. Mayer’s silicone-jar method from the approaches?
5. Page 10. This approach only contemplates application to Superfund projects?
6. Page 10. The proposed method, as described, excludes compounds that are expected to produce greatest concern through bioaccumulation in the food chain, e.g., PCBs. However, bioavailability of the compounds, as reflected by *C_{free}*, is central to the movement of contaminants from the sediment into the food chain. Removing these compounds from inclusion in a more general approach seems unjustified; however, I realize that the method would have to be expanded to include such indirect pathways. Inclusion of this pathway in the method is needed in order for the approach to have meaningful benefit to Superfund.
7. Page 12. The guidance should provide specific information about how a user should determine whether they have a NAPL problem or not. How much NAPL is too much NAPL for the method to perform, overall, at a site? How would a project team go about determining this?
8. Page 13, last sentence. How would a user determine their level of confidence?

9. Page 14, first sentence. Having a benchmark that determines the level above which effects are unacceptable is different than having a benchmark that sets the level below which effects are acceptable. Which do you mean in this case?
10. Page 14, paragraph 1, sentence beginning with "Because". This is a weak argument for justifying the adoption of chronic exposures. You could say that in the majority of contaminated sediment sites populations can be exposed to contaminants for the entire life cycle organism and multiple generations.
11. Page 14, first paragraph. Is the use of 95% of tested species a policy of Superfund?
12. Page 15. If a Superfund site identified a particular benthic species or taxon as the receptor of concern, could the method allow for such an application as an alternative to using the FCV, which potentially considers a much larger group of taxa?
13. Page 15, second to last sentence. There is a log unit of variation around the line. This variation would seem to have implications for the reliability of applying the method. The document should discuss those implications.
14. Page 22. Would be beneficial to include a list of specific sediment parameters that should be measured, and how, as a part of "bulk sediment" analysis the data are to be used in the IWRG method. Also, would be useful to include a scenario in the description where you have widely varying contaminant bioavailability due to differences in OC content, or the nature of the OC present. This scenario will cause challenges in using the method in decision-making and management for large, heterogeneous sites.
15. Page 24 last sentence third paragraph and first sentence fourth paragraph. This description does not provide sufficient guidance. How should a user determine what is small or wide variability?
16. Page 24. Kocs?
17. Page 24, fourth paragraph. Variability in Koc across a site presents a major challenge to applying this method. This paragraph does not provide sufficient guidance to address this problem.
18. Page 25, first full sentence. This guidance is not sufficient. You're telling them to be conservative without specifics for how to do this or a discussion of the implications of doing so.
19. Page 25, second to last sentence. What should be considered too large? The guidance here is not sufficient.
20. Page 28. How does all of this work if the nature of the OC varies across the site?
21. Page 28, last paragraph. These alternative approaches may simplify the arithmetic, but I don't think they simplify the assessment, considering the assumptions involved. What are the implications of using these "shortcuts"?
22. Page 31. The mixture problem is a challenge with the method. Are there any test cases you can refer to that evaluate how robust the method with respect to the assumptions of independent action and additivity? One approach to simplifying the problem would be to recommend screening out contaminants that are not expected to make a meaningful contribution to toxicity.
23. Page 32 and following. The presentation of information and text in Section 5 of the document confused me. I recommend that that the material contained in this section be reorganized with the purpose of providing guidance to the field practitioner on how the described method can be used as a part of a multiple lines-of-evidence approach to support remedial decision making.

24. Page 32, "...and there are a host of issues that could arise with the passive sampling technique." This statement and following text is not organized or presented in a way that contributes to guidance for field application. More text and explanation is needed here.
25. Page 33, section 5.2, first sentence. This sentence says that toxicity tests provide toxicity data. It's an odd first sentence.
26. Page 37, first paragraph. Are you recommending that a Superfund project should develop water-only chemical data if they don't have a full dose-response curve? I question the practicality of this as a recommendation.
27. Page 37, second paragraph. How many cases are there of Superfund projects using TIE? My sense is that there would be a very small number of cases, separate from R&D studies conducted with Superfund site sediments. I question the practicality of this recommendation. I think you need a stronger argument here to justify the additional effort involved.
28. Page 37, third paragraph and elsewhere. I question the practicality of the brief recommendations for additional sampling. This would be more implementable at small, spatially simple sites. However, the general call for collecting more data begs the question of how much data is needed in order to apply the method. The existing text does not treat this issue of data adequacy in a sufficient way. There are significant pressures at play at Superfund sites regarding data collection. Recommendations for additional data should be supported by guidance that provides the means for justifying the effort and expense of collecting the data.
29. Page 38, Section 5.5, first paragraph. The discussion of in situ vs. ex situ approaches is not sufficient to provide guidance to Superfund project teams. In addition to uncertainties and complexities, the purpose or intention to be served by the data should determine whether in situ or ex situ approaches would be most applicable. This fact and others are missing from the discussion. I don't see how a Superfund project manager would make use of this brief discussion. I think it would be beneficial to discuss some specific case applications of in situ and ex situ methods to illustrate the main points and distinctions.
30. Page 39, second paragraph. The discussion of multiple replication schemes is too open-ended to be useful guidance. In fact, I think project managers will find this discussion off-putting, given the implied costs. Need to better define how the information about variance would be used for the types of decisions Superfund teams are supporting.
31. Page 40. The information provided in the Appendix introduces these issues but doesn't provide guidance that is specific or detailed enough for Superfund project teams to apply. What specific information would a project team need to draw upon to reach a decision regarding these issues and how to deal with them?